

A randomness test for functional panels

Piotr Kokoszka ^{*1}, Matthew Reimherr ^{†2}, and Nikolas Wölfing ^{‡3}

¹Colorado State University

²Pennsylvania State University

³Centre for European Economic Research, ZEW Mannheim

Abstract

Functional panels are collections of functional time series, and arise often in the study of high frequency multivariate data. We develop a portmanteau style test to determine if the cross-sections of such a panel are independent and identically distributed. Our framework allows the number of functional projections and/or the number of time series to grow with the sample size. A large sample justification is based on a new central limit theorem for random vectors of increasing dimension. With a proper normalization, the limit is standard normal, potentially making this result easily applicable in other FDA context in which projections on a subspace of increasing dimension are used. The test is shown to have correct size and excellent power using simulated panels whose random structure mimics the realistic dependence encountered in real panel data. It is expected to find application in climatology, finance, ecology, economics, and geophysics. We apply it to Southern Pacific sea surface temperature data, precipitation patterns in the South-West United States, and temperature curves in Germany.

1 Introduction

We define a *functional panel* as a stochastic process of the form

$$(1) \quad \mathbf{X}_n(t) = [X_{1,n}(t), \dots, X_{I,n}(t)]^\top, \quad 1 \leq n \leq N,$$

where each $X_{i,n}$ is a function of time t . The dimension I can increase with the series length N , with examples discussed below. For the applications that motivate the present research,

^{*}Piotr.Kokoszka@colostate.edu

[†]*Corresponding author*, mreimherr@psu.edu

[‡]Woelfing@zew.de, funding provided by the Helmholtz Association through the Helmholtz Alliance Energy-Trans is gratefully acknowledged.

it is enough to think of the $X_{i,n}$ as curves defined on the same time interval, but in principle, functions on more general domains, e.g., volumes or surfaces, can be considered. The discrete time index n refers to a unit like a day, week or year. The index t is the continuous time argument of the function $X_{i,n}$. The index i refers to the i^{th} time series in the panel. This paper develops a test of the null hypothesis

H_0 : the random elements \mathbf{X}_n , $1 \leq n \leq N$, are independent and identically distributed.

Our test is designed to detect serial dependence, and we assume stationarity across n even under the alternative.

To illustrate the functional panel concept, Figure 1 shows four curves, $I = 4$, for a fixed n . The index n refers to years, and the four curves describe the sea surface temperature in four regions used to measure the El Niño climatic phenomenon. Figure 2 shows another example, now with i fixed. The data point $X_{i,n}(t)$ is the log-precipitation at location i on day t of year n . The construction of this series is explained in detail in Section 3. Data structures of this type are very common in climate studies; $X_{i,n}(t)$ can be total precipitation or maximum temperature on day t , $1 \leq t \leq 365$, of year n at location i in some region. In such climate applications, I is comparable to N because records often start at the end of the 19th or towards the middle of the 20th century, thus, they are about 60 to 120 years long ($N \approx 60$ to 120), and there are several dozen measurement stations in a region ($I \approx 40$ to 120). (The United States Historical Climatology Network – USHCN – contains weather data collected at 1,218 stations across the 48 contiguous states, starting from ca. 1900.)

Climate data do not exhaust possible applications. Intraday financial data typically come in panels. For example, $X_{i,n}(t)$ can be the exchange rate (against the US dollar) of currency i , $1 \leq i \leq I$, at minute t of the n^{th} trading day. Panels of exchange rates contain information on the intraday strength of the US dollar. Corporate bond yield curves are large panels because a bond portfolio includes hundreds of companies, $I \sim 10^3$; in economic studies, government bond yields curves form small panels because only a few countries are considered to assess risk in a region, see, e.g., Härdle and Majer [19]. At the intersection of climate and financial panels, Härdle and Osipienko [18] use a functional panel framework in which i refers to a spatial location, and the interest lies in pricing a financial derivative product whose value depends on the weather at location i . Modeling electricity data involves functional panels indexed by regions or power companies with the daily index n , see Liebl [32] for an overview. Daily pollution (particulate, oxide or ozone) curves at several locations within a city form a functional panel of moderate size.

In these examples, the dependence between the $X_{i,n}$, $1 \leq i \leq I$, for fixed n , is strong, and, generally, the temporal dependence, indexed by n , cannot be neglected. In specific applications, this dependence is modeled by deterministic trends or periodic functions (climate data) or by common factors (financial data). To validate a model, it is usual to verify that residual curves computed in some manner form a random sample. (See, e.g., Kowal et al. [29] for a model applied to functional panels of government yield curves and neurological measurements with an explicit residual iid assumption.) It is thus important to develop a

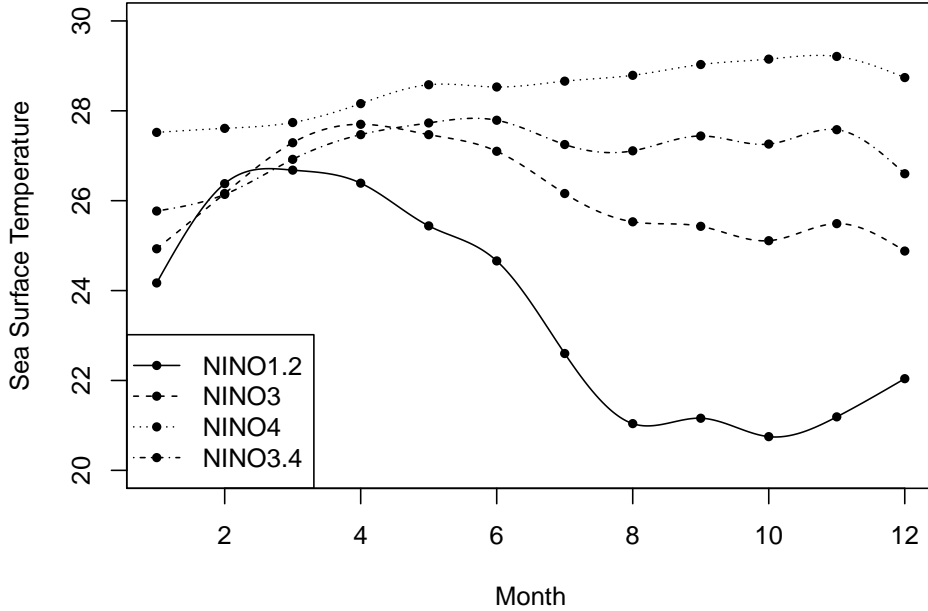


Figure 1: Sea Surface Temperature curves of El Niño regions in 2012

test of randomness, i.e., to test the null hypothesis H_0 stated above. Such a test could be viewed as analogous to tests of randomness which are crucial in time series analysis, see, e.g., Section 1.6 of Brockwell and Davis [5]. They can be applied to original or transformed data, or to model residuals. The purpose of this paper is to develop a suitable test for functional panels. Before discussing our approach, we provide some historical background. Our methodology builds on the well-established paradigm of testing for randomness in time series which can be traced back to the work of Box and Pierce [3], which was followed by a number of influential contributions including Chitturi [8], Hosking [25], Ljung and Box [33] and McLeod [35]. These tests use as a starting point the asymptotic distribution of the sample autocorrelations of a white noise: the $\hat{\rho}_h$ are approximately independent normal random variables with mean zero and variance $1/N$, where N is the sample size. Therefore, $N \sum_{h=1}^H \hat{\rho}_h^2$ is approximately chi-square with H degrees of freedom. This research is now reported in textbook expositions including Brockwell and Davis [4], Li [31], and Lütkepohl [34]. More recent contributions include Fisher and Gallagher [12] and Peña and Rodriguez [40]. For a single functional time series, a randomness test was derived by Gabrys and Kokoszka [15] and elaborated on by Horváth et al. [23] and Jiofack and Nkiet [27]. In the context of *scalar* panel data, the only work we are aware of is Fu et al. [14] who define

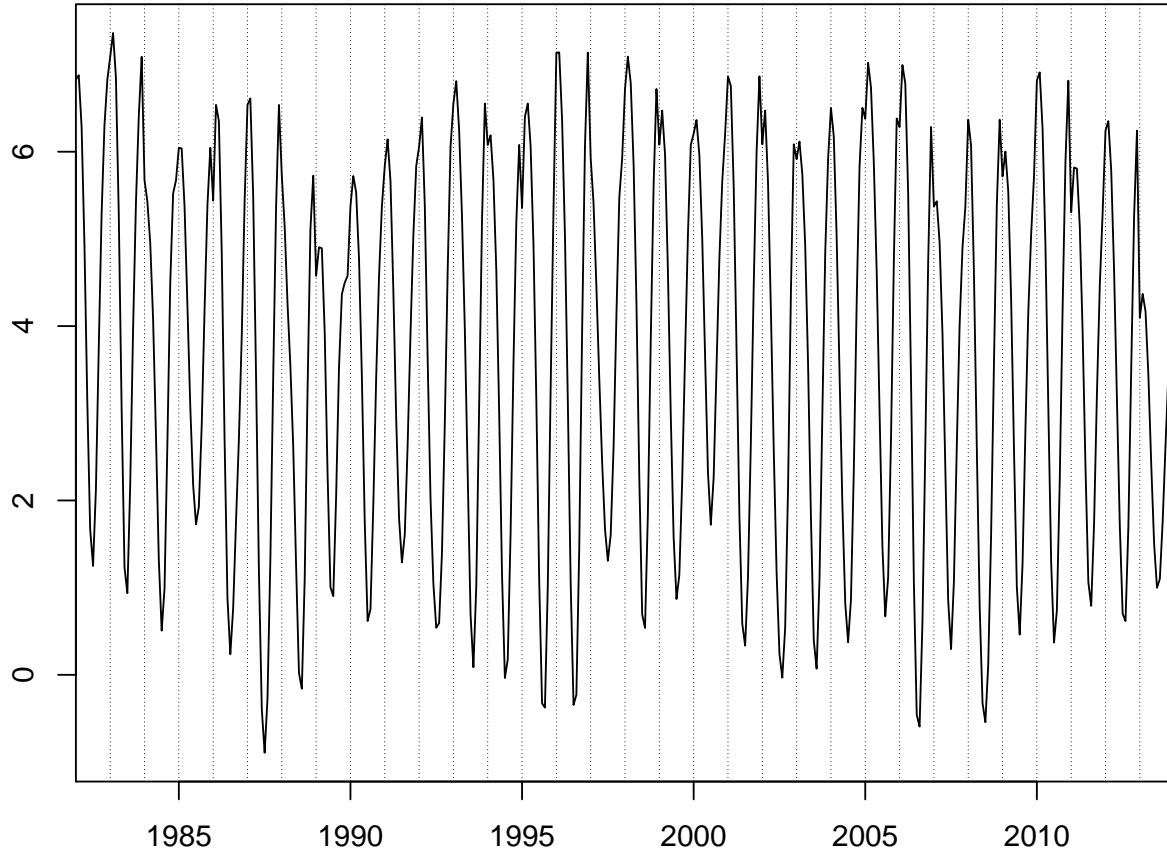


Figure 2: Smoothed log-precipitation, Santa Cruz, California, 1982 to 2013.

residual autocorrelations in the autoregressive panel model of Hjellvik and Tjøstheim [20] as

$$\hat{r}_h = \frac{\sum_{i=1}^I \sum_{n=1}^{N-h} \hat{\varepsilon}_{i,n+h} \hat{\varepsilon}_{i,n}}{\sum_{i=1}^I \sum_{n=1}^{N-h} \hat{\varepsilon}_{i,n}^2},$$

where the $\hat{\varepsilon}_{i,n}$ are appropriately defined residuals. In their asymptotic setting, the number of temporal points, N , is fixed, and the number of time series, I , increases to infinity. They show that for any fixed H , the vector $[\hat{r}_1, \dots, \hat{r}_H]^\top$ is asymptotically normal with the asymptotic covariance matrix that can be estimated. By constructing a suitable quadratic form, they derive a portmanteau test statistic whose asymptotic distribution is χ_H^2 . There is at present no randomness test for functional panels, and it is our objective to derive a practically useable test which is supported by asymptotic arguments. Hsiao [26] provides an excellent account of the methodology for scalar panel data.

We reduce the dimension of the functions $X_{i,n}$ by using projections on functional principal components. Denote the number of such projections for the i^{th} series in the panel by $p(i)$. The total number of scalar time series we must consider is thus $p = \sum_{i=1}^I p(i)$. For climate applications discussed above, p can approach several hundred. It is thus natural to consider

asymptotics with p increasing to infinity with N . Our theory applies to cases of fixed $p(i)$ and increasing I , fixed I and increasing $p(i)$, or both increasing with N . In this framework, we show that it is possible to construct a test statistic that is asymptotically standard normal. The work of Fu et al. [14] can be viewed as considering $p(i) = 1$, N fixed, and $I \rightarrow \infty$. Our setting is thus quite different, and requires a new asymptotic framework. Despite some theoretical complexity, our approach leads to a test whose asymptotic null distribution is standard normal, and which can be algorithmically implemented. It is therefore hoped that our work will find application in the analysis of functional panels of the type specified above. In particular, it could motivate the development of suitable change point tests that target more specific alternatives, Aston and Kirch [1] and Zhang et al. [45] consider the case of $I = 1$.

Since the goal of our methodology is to check the independence assumption, our tests are based on the usual functional principal component scores; functional principal components form the optimal basis under the null hypothesis. They have an established place in FDA research with readily available `R` and `matlab` implementations. In principle, other basis systems could be used, especially those custom-developed for time series of functions, see, e.g., Hörmann et al. [22] and Panaretos and Tavakoli [36], or even those going beyond linear dimension reduction, see Li and Song [30]. In each of these cases, our general approach could be applied to the resulting scores, but new asymptotic justifications and numerical implementations would have to be developed.

The remainder of the paper is organized as follows. In Section 2.1, we formalize the asymptotic framework, derive the test statistic and establish its asymptotic normality. Section 2.3 describes the practical implementation of the test procedure in algorithmic steps. Section 3 illustrates the application of the test on three climate data sets which form functional panels: sea surface temperatures in the pacific ocean, US regional precipitation data and temperature curves in Germany. Section 3.2 further examines the finite sample performance of our test by applying it to simulated data which resemble the above mentioned data sets. The proofs of the asymptotic results are presented in Section A and in the supplemental material. In addition to these mathematical calculations, the supplemental material contains a zipped folder containing the complete `R` code, a corresponding `README` file and the data sets.

2 Testing procedure

2.1 Assumptions and large sample results

We assume that all functions have been rescaled so that their domain is the unit interval $[0, 1]$. We also assume that they have mean zero: $E X_{i,n}(t) = 0$ for almost all $t \in [0, 1]$. In practice, the mean is removed by subtracting the sample mean, see Section 2.3, so that the functional time series forming the panel each have sample mean zero. Subtracting the sample mean introduces additional terms of the order $O_P(N^{-1})$, and so does not affect the

limiting distribution.

Denote by $L^2 = L^2([0, 1])$ the Hilbert space of square integrable functions with the usual inner product $\langle \cdot, \cdot \rangle$ and the norm $\| \cdot \|$ it generates. The assumptions and the definition of the test statistic involves the Kronecker product, and its properties are heavily used in the proofs. Readers are referred to Graham [16] for a very useful exposition. In the context of matrices and vectors, we take \otimes to be the usual Kronecker product. Between two functions or operators x and y , we take $x \otimes y$ to be the operator $\langle x, \cdot \rangle y$. We will often not distinguish notationally between the cases, as it will always be clear from the context which we mean. Further details will be provided as needed. By $|\cdot|$ we denote the Euclidean norm of a vector.

Our first assumption states the functions forming the panel are in L^2 and have uniformly bounded fourth moments.

Assumption 1. *Assume that $\{\mathbf{X}_n\}$ is a zero mean sequence of random functional vectors taking values in $\{L^2\}^I$. Furthermore, assume that there exists a constant M such that*

$$\mathbb{E} \|X_{i,n}\|^4 \leq M < \infty, \quad i = 1, \dots, I, \quad n = 1, \dots, N.$$

Our second assumption connects the rate of growth of I and the $p(i)$ to the rate of decay of the gaps between the eigenvalues of individual series and the rate of decay of the eigenvalues of the whole panel. Assumptions of this type go back at least to the work of Dauxois et al. [11]. To the best of our knowledge, only the case of a single functional series or sample, possibly with explanatory variables or functions, has been considered, see Cai and Hall [6], Crambes et al. [10], Fremdt et al. [13], Hall et al. [17], and Paul and Peng [37], among many others. The complexity of our Assumption 2 is due to the panel structure of the data. To formulate it, define

$$\mathbf{X}_{in} = [X_{1in}, \dots, X_{p(i)in}]^\top, \quad X_{jin} = \langle X_{i,n}, v_{i,j} \rangle$$

and column vectors $\mathbf{X}_n^* = [\mathbf{X}_{1n}^\top, \dots, \mathbf{X}_{In}^\top]^\top$ of length $p_N := p = \sum_{i=1}^I p(i)$. The panel $\{\mathbf{X}_n^*\}$ is thus an approximation of dimension p_N to the functional panel $\{\mathbf{X}_n\}$ given by (1). Let

$$\mathbf{C}_{0,N} = \mathbb{E} [\mathbf{X}_n^* \mathbf{X}_n^{*\top}]$$

be the $p_N \times p_N$ covariance matrix whose eigenvalues are $\gamma_1 \geq \dots \geq \gamma_{p_N}$. Denote by $\lambda_{i,1} > \lambda_{i,2} > \dots$ the eigenvalues of the covariance operator $\mathbb{E}[X_{i,1} \otimes X_{i,1}]$ and define

$$\Gamma_N = \sum_{i,i'=1}^I \sum_{j=1}^{p(i)} \sum_{j'=1}^{p(i')} (\alpha_{i,j}^{-1} + \alpha_{i',j'}^{-1})^2,$$

where $\alpha_{i,1} = \lambda_{i,1} - \lambda_{i,2}$ and for $j \geq 2$, $\alpha_{i,j} = \min\{\lambda_{i,j-1} - \lambda_{i,j}, \lambda_{i,j} - \lambda_{i,j+1}\}$.

Assumption 2. *Assume that the sequence p_N is such that $p_N \rightarrow \infty$ and*

$$N^{-1/2} p_N^{-1} \gamma_{p_N}^{-3} I^3 \Gamma_N^{1/2} \rightarrow 0.$$

(The number of panels, I , can either stay fixed or tend to infinity.)

Assumption 2 has the following interpretation. The first two terms, $N^{-1/2}p_N^{-1}$, indicate the rate at which information accumulates as $N \rightarrow \infty$, while the third and fourth terms, $\gamma_{p_N}^{-3}I^3$, indicate the rate at which the panel structure detracts information (with γ_{p_N} governing the correlation between series). The last term $\Gamma_N^{1/2}$ incorporates the spacing of the eigenvalues and is common in asymptotics with an increasing number of projections. A more readily interpretable form of this assumption is stated in Section 2.2 for the case of a single time series ($I = 1$). We do not impose any specific dependence structure, and prefer to use a general, admittedly rather technical, Assumption 2. An alternative approach would be to impose some temporal dependence structure, e.g., as in Jirak [28], and establish analogous results under such assumptions. Instead, we give a brief example to help shed further light on Assumption 2.

Assume that each element of the panel has the same covariance operator so that $\lambda_{i,j} \equiv \lambda_j$ and $\alpha_{i,j} \equiv \alpha_j$ for all i and j . In this case, it makes sense to also assume that $p_i \equiv p$ so that $p_N = Ip$. Collect the λ_j into a diagonal matrix Λ . Furthermore, assume that the panels are independent so $\mathbf{C}_{0,N} = \Lambda \otimes \mathbf{I}_{I \times I}$, where \otimes denotes the Kronecker product and $\mathbf{I}_{I \times I}$ the $I \times I$ identity. This then implies that $\gamma_{p_N} = \lambda_p$.

We now assume explicitly that $\lambda_j = j^{-\alpha}$ and $\lambda_j - \lambda_{j+1} = j^{-\alpha-1}$. This implies that

$$\begin{aligned} \Gamma_N &\leq I^2 \sum_j \sum_{j'} [j^{\alpha+1} + (j')^{\alpha+1}]^2 \\ &= I^2 \left[2p \sum_{j=1}^p j^{2\alpha+2} + 2 \left(\sum_{j=1}^p j^{\alpha+1} \right)^2 \right] \\ &\approx I^2 [2p^{2\alpha+4}(2\alpha+3)^{-1} + 2p^{2\alpha+4}(\alpha+2)^{-2}] \sim I^2 p^{2\alpha+4}. \end{aligned}$$

Here \approx means the limit of their ratio tends to 1, while \sim means the limit of their ratio is a finite nonzero constant. So then we have

$$N^{-1/2}p_N^{-1}\gamma_{p_N}^{-3}I^3\Gamma_N^{1/2} \sim N^{-1/2}p^{-1}I^{-1}p^{3\alpha}I^3Ip^{\alpha+2} = N^{-1/2}p^{4\alpha+1}I^2.$$

The parameter α is usually viewed as the smoothness of the $X_{i,n}$ processes. We can see that for rougher processes, we can actually take larger panels and more principal components, since α will be smaller in these cases. For example, $\alpha = 2$ for Brownian motion. The same calculations will show that in the single panel case of Section 2.2, the rate becomes

$$\frac{N^{-1/2} \sum_{j=1}^p \alpha_j^{-1}}{\lambda_p^2} \sim N^{-1/2} p^{3\alpha+2}.$$

Since it must be the case that $\alpha > 1$, we can see the price we pay for the lack of structure in the panel as $3\alpha + 2 < 4\alpha + 1$. This price increases for smoother processes, i.e., larger α .

We now proceed to define the test statistic. Let $\hat{v}_{i,j}$ be the j^{th} estimated functional principal component (EFPC) of the i^{th} functional time series, see, e.g., Chapter 3 of Horváth and Kokoszka [24]. Set

$$\hat{\mathbf{X}}_{in} = [\hat{X}_{1in}, \dots, \hat{X}_{p(i)in}]^\top, \quad \hat{X}_{jin} = \langle X_{i,n}, \hat{v}_{i,j} \rangle.$$

Next, we form column vectors of length $p_N := p = \sum_{i=1}^I p(i)$ given by

$$\widehat{\mathbf{X}}_n = [\widehat{\mathbf{X}}_{1n}^\top, \dots, \widehat{\mathbf{X}}_{In}^\top]^\top.$$

To form a portmanteau test statistic using the $\widehat{\mathbf{X}}_n$, we introduce

$$\widehat{\mathbf{V}}_h = N^{-1} \sum_{n=1}^{N-h} \widehat{\mathbf{X}}_n \otimes \widehat{\mathbf{X}}_{n+h}; \quad \widehat{\mathbf{C}}_0 = N^{-1} \sum_{n=1}^N \widehat{\mathbf{X}}_n \widehat{\mathbf{X}}_n^\top.$$

Observe that $\widehat{\mathbf{V}}_h$ is a column vector of length p_N^2 and $\widehat{\mathbf{C}}_0 \otimes \widehat{\mathbf{C}}_0$ is a $p_N^2 \times p_N^2$ symmetric matrix. The test statistic is defined by

$$\widehat{Q}_N = N \sum_{h=1}^H \widehat{\mathbf{V}}_h^\top (\widehat{\mathbf{C}}_0 \otimes \widehat{\mathbf{C}}_0)^{-1} \widehat{\mathbf{V}}_h.$$

The summation limit H plays the same role as the maximal number of lags in the usual Box–Pierce–Ljung type statistics. It is fixed in the asymptotic theory.

Our first result states that \widehat{Q}_N is asymptotically normal under H_0 , i.e., when the data are iid.

Theorem 1. *If Assumptions 1 and 2 hold, then under H_0 ,*

$$\frac{\widehat{Q}_N - p_N^2 H}{p_N \sqrt{2H}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Theorem 1 is proven in Appendix A. The proof involves a sequence of vectors of projections of increasing dimension. In Cardot et al. [7] and Horváth et al. [23], this problem is avoided by making extensive use of the Prokhorov–Levy metric. However, such a technique is limited due to the difficulty of incorporating dimension into any Berry–Esseen type convergence result, which typically rely on highly complex smoothing arguments. Furthermore, such an approach typically does not yield results as sharp as proving the CLT directly due to the way they depend on dimension. In contrast, our approach adds no additional assumptions beyond those needed to replace the estimated eigenvalues and eigenfunctions with their theoretical counterparts. We therefore view the following theorem, which establishes the asymptotic normality of general quadratic forms based on autocorrelations, as an important contribution of this paper.

Theorem 2. *Let $\{\mathbf{Z}_{n,N}, 1 \leq n \leq N\}$ be an array of random vectors with $\mathbf{Z}_{n,N} \in \mathbb{R}^{p_N}$. For each N , assume that $\mathbf{Z}_{1,N}, \dots, \mathbf{Z}_{N,N}$ are iid and that*

$$(2) \quad \mathbb{E}[\mathbf{Z}_{1,N}] = \mathbf{0} \quad \text{and} \quad \mathbb{E}[\mathbf{Z}_{1,N} \mathbf{Z}_{1,N}^\top] = \mathbf{I}_{p_N}.$$

If, as $N \rightarrow \infty$,

$$(3) \quad p_N \rightarrow \infty, \quad p_N N^{-2/3} \rightarrow 0, \quad \text{and} \quad N^{-1/2} \mathbb{E} |\mathbf{Z}_{1,N}|^4 \rightarrow 0,$$

then for H fixed

$$\frac{N^{-1} \sum_{h=1}^H \left| \sum_{n=h+1}^N \mathbf{Z}_{n-h,N} \otimes \mathbf{Z}_{n,N} \right|^2 - p_N^2 H}{p_N \sqrt{2H}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

The proof of Theorem 2 is presented in the supplemental material. Limit results for random vectors with the dimension increasing with the sample size appear in the asymptotic theory for empirical likelihood, see, e.g., Hjort et al. [21] and Peng and Schick [39]. The Central Limit Theorem established by Peng and Schick [38] is motivated by such theory. Using the notation of Theorem 2, a corollary to their main result can be stated as

$$\frac{N^{-1} \left| \sum_{n=1}^N \mathbf{Z}_{n,N} \right|^2 - p_N}{\sqrt{2p_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Their focus is not on the lagged Kronecker products, but on the case where $E[\mathbf{Z}_{1,N} \mathbf{Z}_{1,N}^\top]$ is a general covariance matrix (i.e., not the identity), and the centering is with respect to its trace. Other results on CLT convergence rates which incorporate dimension are given in Senatov [42].

We conclude this section with a general framework under which the test rejects the null. If the sequence is stationary and weakly dependent and at least one element of the panel exhibits nonzero correlation with another element (at some lagged time index), then the test will reject with power approaching one. As a specific assumption for stationarity and weak dependence we use the concept of L^p - m -approximability, see Horváth and Kokoszka [24], Chapter 16.

Assumption 3. (*Alternative Hypothesis*) Assume that $\{(X_{i,1}, \dots, X_{i,N})\}$ is a stationary L^4 - m approximable sequence (for each i) and that there exists a nonempty subset of indices $\mathcal{H}^* \subset \{1, \dots, H\}$, and for each $h \in \mathcal{H}^*$ a nonempty subset of pairs of indices $\mathcal{I}_h^* \subset \{1, \dots, I\}^2$ such that and $j \in \{1, \dots, I\}$ such that

$$(E[\langle X_n, v_i \rangle \langle X_{n+h}, v_j \rangle])^2 \geq R > 0,$$

for all $h \in \mathcal{H}^*$ and $(i, j) \in \mathcal{I}_h^*$.

Theorem 3. If Assumptions 1, 2, and 3 hold, then

$$\frac{\hat{Q}_N - p_N^2 H}{p_N \sqrt{2H}} \gtrsim \gamma_1^{-2} R N (1 + o_P(1)) \sum_{h \in \mathcal{H}^*} |\mathcal{I}_h^*| \xrightarrow{P} \infty.$$

Theorem 3 is proven in Appendix A

In the next section, we consider the case of a single series to illustrate our assumptions. We emphasize that H is assumed to be fixed, but can be arbitrarily large. Asymptotic under H diverging to infinity (for a single series) were investigated by Horváth et al. [23]. It should,

in principle, be possible to let the number of panel series, I , the number of projections p and the maximal lag H tend simultaneously to infinity, but we do not develop such a more complex theory here. We thus stay within the framework of traditional time series analysis where asymptotics are derived for a finite number of lags, see, e.g., Chapter 7 of Brockwell and Davis [4].

2.2 Case of $I = 1$ (a single functional series)

To provide a more tangible intuition behind the form of the statistic \widehat{Q}_N , we discuss the simpler scenario where we only have one time series, i.e., $I = 1$. Define

$$\Delta_{N,h} = N^{-1/2} \sum_{n=1}^{N-h} X_n \otimes X_{n+h}.$$

The autocovariance operator $\Delta_{N,h}$ is Hilbert–Schmidt. Recall that Hilbert–Schmidt operators form a separable Hilbert space with the inner product

$$(4) \quad \langle \Psi_1, \Psi_2 \rangle_{\mathcal{S}} = \sum_k \langle \Psi_1(e_k), \Psi_2(e_k) \rangle,$$

where $\{e_k\}$ is any orthonormal basis, see, e.g., Chapter 2 of Horváth and Kokoszka [24]. A direct application of this definition with the functions \hat{v}_k (extended to a complete system) shows that

$$\langle \Delta_{N,h}, \hat{v}_j \otimes \hat{v}_{j'} \rangle_{\mathcal{S}} = N^{-1/2} \sum_{n=1}^{N-h} \widehat{X}_{jn} \widehat{X}_{j',n+h}.$$

Therefore, by Lemma 7.1 of Horváth and Kokoszka [24], the statistic \widehat{Q}_N can be expressed as

$$\widehat{Q}_N = \sum_{h=1}^H \sum_{j,j'=1}^{p_N} \frac{\langle \Delta_{N,h}, \hat{v}_j \otimes \hat{v}_{j'} \rangle_{\mathcal{S}}^2}{\hat{\lambda}_j \hat{\lambda}_{j'}}.$$

The summands are the squares of the sample cross-correlations of all projections under consideration. These are added over all projections and all lags up to lag H .

In the case of a single time series, Assumption 2 can be replaced by a more interpretable assumption:

Assumption 4. *We assume that the sequence p_N is nondecreasing, $p_N \rightarrow \infty$, and satisfies*

$$\frac{N^{-1/2} \sum_{j=1}^{p_N} \alpha_j^{-1}}{\lambda_{p_N}^2} \rightarrow 0,$$

where $\alpha_1 = \lambda_1 - \lambda_2$ and for $j \geq 2$, $\alpha_j = \min\{\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1}\}$.

Assumption 4 quantifies the intuition that p should increase to infinity at a rate slower than N , depending on the rate of decay of the eigenvalues λ_j and the gaps between them.

Direct verification shows that if the λ_j decay exponentially fast, then p must increase at a rate slower than $\ln N$. If the λ_j decay like a power function, Assumption 4 will hold if $\ln(p)/\ln(N) \rightarrow 0$. The proof of Theorem 1 for $I = 1$ is presented in the supplemental material. It is less abstract than the general proof in Appendix A; its study may facilitate the understanding of the general case.

2.3 Details of implementation

In this section, we provide a step by step description of the testing procedure. All steps listed below can be easily implemented in **R** or **Matlab** using basic routines and the functional principal component tool box (see Ramsay et al. [41]). Supplemental material contains a ready to use **R** function implementing the test. For step 6, we provide two alternative but equivalent procedures where the first might be more intuitive while the second is computationally more efficient. The finite sample bias correction in step 7 follows from an extension of the arguments of Ljung and Box [33]. It is clearly asymptotically negligible.

1. Center each functional time series, i.e., compute

$$X_{i,n}^c(t) = X_{i,n}(t) - \hat{\mu}_i(t), \quad \hat{\mu}_i(t) = N^{-1} \sum_{n=1}^N X_{i,n}(t).$$

2. Calculate the eigenfunctions $\hat{v}_{i,j}$ and the eigenvalues $\hat{\lambda}_{i,j}$ of the empirical covariance operator defined by

$$\hat{C}_i(x) = \frac{1}{N} \sum_{n=1}^N \langle X_{i,n}^c, x \rangle X_{i,n}^c.$$

(This step is implemented as `pca.fd` in **R** and as `pca_fd` in **Matlab**. Both functions, by default, center their arguments as in step 1.)

3. For each $1 \leq i \leq I$, determine $p(i)$ as the smallest k for which

$$\frac{\sum_{j=1}^k \hat{\lambda}_{i,j}}{\sum_{j=1}^N \hat{\lambda}_{i,j}} > 0.85.$$

4. Construct the vectors of scores

$$\hat{\mathbf{X}}_{in} = [\hat{X}_{1in}, \dots, \hat{X}_{p(i)in}]^\top, \quad \hat{X}_{jin} = \langle X_{i,n}, \hat{v}_{i,j} \rangle.$$

and the vectors

$$\hat{\mathbf{X}}_n = [\hat{\mathbf{X}}_{1n}^\top, \dots, \hat{\mathbf{X}}_{In}^\top]^\top.$$

Using the vectors $\{\hat{\mathbf{X}}_n\}$, generate the $N \times p_N$ matrix

$$\hat{\mathbf{X}} = [\hat{\mathbf{X}}_1, \dots, \hat{\mathbf{X}}_N]^\top$$

and calculate the empirical covariance matrix

$$\hat{\mathbf{C}}_0 = N^{-1} \hat{\mathbf{X}}^\top \hat{\mathbf{X}}.$$

5. Calculate the spectral decomposition

$$\hat{\mathbf{C}}_0 = \mathbf{U}\mathbf{D}\mathbf{U}^\top,$$

where \mathbf{D} is the diagonal matrix of eigenvalues and \mathbf{U} the matrix of eigenvectors. Let d_1, \dots, d_{p_N} be the eigenvalues of $\hat{\mathbf{C}}_0$. Choose a cutoff point q defined as the smallest integer for which

$$\frac{\sum_{i=1}^q d_i}{\sum_{i=1}^{p_N} d_i} \geq 0.85.$$

Set $\mathbf{D}^{-1}[i, i] = d_i^{-1}$ if $1 \leq i \leq q$ and zero otherwise. Calculate the generalized inverse

$$\hat{\mathbf{C}}_0^{-1} = \mathbf{U}\mathbf{D}^{-1}\mathbf{U}^\top.$$

(Note that $\hat{\mathbf{C}}_0$ might be singular, e.g., when $p_N > N$.)

6. Using the vectors $\{\hat{\mathbf{X}}_n\}$ from step 4, compute the terms

$$\hat{\mathbf{V}}_h = N^{-1} \sum_{n=1}^{N-h} \hat{\mathbf{X}}_n \otimes \hat{\mathbf{X}}_{n+h}$$

for $1 \leq h \leq H$, where H is chosen by the user. (A discussion of this issue is presented at the end of Section 3.2.)

The test statistic

$$\hat{Q}_N = N \sum_{h=1}^H \hat{\mathbf{V}}_h^\top (\hat{\mathbf{C}}_0 \otimes \hat{\mathbf{C}}_0)^{-1} \hat{\mathbf{V}}_h.$$

can be calculated using $(\hat{\mathbf{C}}_0 \otimes \hat{\mathbf{C}}_0)^{-1} \approx \hat{\mathbf{C}}_0^{-1} \otimes \hat{\mathbf{C}}_0^{-1}$, where $\hat{\mathbf{C}}_0^{-1}$ is the generalized inverse from step 5.

An alternative procedure for step 6 which avoids the Kronecker product and therefore requires less computational resources is the following:

For each $h, 1 \leq h \leq H$, take two submatrices of the matrix $\hat{\mathbf{X}}$ defined in step 4,

$$\begin{aligned} \hat{\mathbf{X}}_{-h} &= [\hat{\mathbf{X}}_1, \dots, \hat{\mathbf{X}}_{N-h}]^\top \\ \hat{\mathbf{X}}_{+h} &= [\hat{\mathbf{X}}_{1+h}, \dots, \hat{\mathbf{X}}_N]^\top \end{aligned}$$

and construct the matrices

$$\hat{\mathbf{M}}_h = N^{-1} (\hat{\mathbf{X}}_{-h}^\top \hat{\mathbf{X}}_{+h})^\top, \quad h = 1, \dots, H.$$

(Note that the vectorized form $\text{vec}(\hat{\mathbf{M}}_h)$ is equivalent to the vectors $\hat{\mathbf{V}}_h$ defined before.)

Calculate the test statistic as

$$\hat{Q}_N = N \sum_{h=1}^H \text{vec}(\hat{\mathbf{M}}_h)^\top \text{vec}\left(\left(\hat{\mathbf{C}}_0^{-1}\right)^\top \hat{\mathbf{M}}_h \hat{\mathbf{C}}_0^{-1}\right),$$

where $\hat{\mathbf{C}}_0^{-1}$ is the generalized inverse from step 5.

7. Reject the null hypothesis at significance level $0 < \alpha < 1$, if

$$\frac{\hat{Q}_N - q^2 H \left(1 - \frac{H+1}{2N}\right)}{q \sqrt{2H \left(1 - \frac{H+1}{2N}\right)}} > \Phi^{-1}(1 - \alpha),$$

where q is determined in step 5 and where $\Phi^{-1}(1 - \alpha)$ is the $(1 - \alpha)$ th quantile of the standard normal distribution.

3 Applications and finite sample performance

As discussed in Section 1, there are many examples of functional panels, with various temporal and cross-sectional dependence structures, and various shapes of the curves. This paper focuses on methodology and theory. It is therefore not possible to present a simulation study which covers the wide range of possibly relevant scenarios. However, rather than considering some ad hoc artificial data generating processes (DGP), we focus on three real data sets taken from climate studies and then simulate panels whose random structure resembles the one of these real data sets closely. Our goal is to evaluate the performance of the test in realistic settings and so to provide additional guidance for its application.

3.1 Application to climate data

We consider three climate data sets with different values of I and N and different levels of noise. Each of them consists of N annual curves at I locations. Before describing these data in more detail, we provide the following summary.

1. **El Niño SST:** $N = 63$, $I = 4$, smooth.
2. **US precipitation:** $N = 113$, $I = 103$, noisy.
3. **German temperature:** $N = 61$, $I = 42$, noisy.

El Niño is a phenomenon of semi-periodic variation of sea surface temperature (SST) in the southern Pacific Ocean. The phenomenon is measured by an index for SST variation in several regions, generally referred to as Niño-1+2, Niño-3, Niño-4, and Niño-3.4, see, e.g., Trenberth [43], and Trenberth and Stepaniak [44]. Data with monthly SST measurements for all four regions from 1950 onwards are available online:

<http://www.cpc.ncep.noaa.gov/data/indices/ersst3b.nino.mth.81-10.ascii>. Panels with I comparable to N arise in regional climate studies. As an example, we use monthly precipitation data from the United States Historical Climatology Network for all stations in California, Arizona, and New Mexico, a region known for reoccurring precipitation deficits. Screening for completeness yields a panel of $I = 103$ stations from 1901 to 2013, $N = 113$. Increasing all records by 0.01 inch (which is the smallest discrete unit of measurement for

these data) allows us to use a log-transformation which leads to approximately normal data. (Any invertible transformation preserves the iid property stated as H_0 .) As our last example, we consider daily temperature data at $I = 42$ weather stations in Germany over $N = 61$ years (accessible online at: www.dwd.de).

To remove long term trends that would lead to a rejection, all data were detrended by fitting a least squares line to observations at each location and each month. This operation had a visible effect only on the German temperature data, as illustrated in Figure 3. Arguably, more sophisticated methods of long term behavior modeling could be used, methods, where testing the residuals for independence might be crucial to conclude on the model validity. However, our objective is not a climatological study, but merely an illustration of statistical methodology which can be applied to residual curves obtained from more complex models.

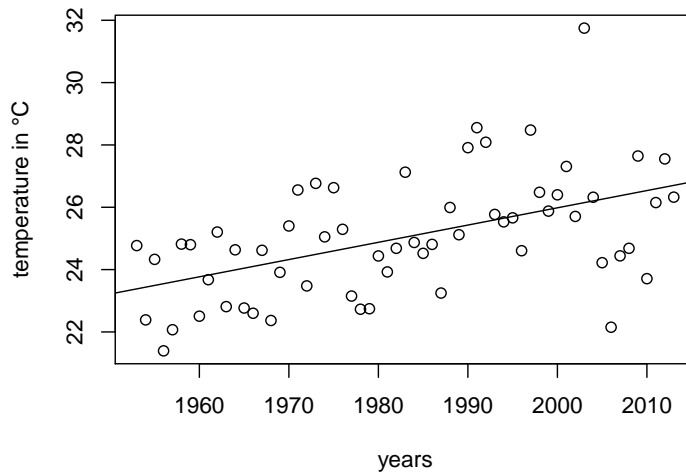


Figure 3: Monthly averages of maximum daily temperature in the month of August, 1953 to 2013, in Mannheim, Germany. Note the very high record of almost 32°C during the European heat wave in Summer 2003.

The El Niño indices are already smoothed regional averages and do not require further smoothing. We therefore expand the monthly measurements using a B-spline basis of order four with one knot placed at each month, the expansion closely matches the data points. Figure 1 shows the spline-expansion for the year 2012, while Figure 4 exhibits the SST measurements in each region over the whole span of the sample. The data from individual weather stations in the US and in Germany are noisy. Due to pronounced annual periodicity, we expand each annual curve using a Fourier series with 25 basis functions and apply the harmonic acceleration roughness penalty. An example for the US smoothed, log-transformed precipitation data is given in Figure 2. The smoothness parameter was chosen to minimize the standard generalized cross-validation criterion. Details of the procedure are described in Section 5.3 of Ramsay et al. [41].

We apply the test to the detrended and smoothed curves. As described in Section 2.3, Step 3, the numbers $p(i)$ are determined separately for each time series i such that at least 85% of variance within the time series is captured by the $p(i)$ principal components. For the El Niño SST curves, this requires two principal components for each region. For the log-precipitation and the temperature data, the same criterion selects three principal components for each station. The fact that all time series within a panel are assigned the same number of principal components underlines the fact that the climate measures we use have structurally similar data generating processes. This, however, does not necessarily need to be the case for different panels. Our testing procedure accounts for such a possibility by choosing $p(i)$ individually for each i . Additional simulations show that the test is very robust to the choice of the cut-off criterion of this first dimension reduction. The second principal component analysis described in Step 5, Section 2.3, determines the number of dimensions q for each sample that are finally used to derive our test statistic. The 85% of variance criterion selects $q = 2$ for the El Niño data with $I = 4$. It selects $q = 3$ for the German temperature data with $I = 42$, and $q = 37$ for the US precipitation data with $I = 103$. The fact that the second dimension reduction for the German temperature data selects just 3 principal components for the whole panel can be attributed to the rather dense geographical coverage and thus large homogeneity of the cross-sectional temperature curves in this sample.

Table 1: Test results (normalized \hat{Q}_N and p -values) for the three data sets: El Niño SST curves with $I = 4$, $N = 63$; US log-precipitation data with $I = 103$, $N = 113$; German temperature data with $I = 42$, $N = 61$.

	El Niño SST		US precip.		German temp.	
	stat.	p -value	stat.	p -value	stat.	p -value
$H = 1$	13.483	< 0.001	1.704	0.044	-0.701	0.758
$H = 2$	11.778	< 0.001	2.080	0.019	-1.000	0.841
$H = 3$	10.282	< 0.001	2.827	0.002	-0.138	0.555
$H = 4$	8.973	< 0.001	2.587	0.005	0.193	0.424
$H = 5$	7.617	< 0.001	2.714	0.003	-0.349	0.636
$H = 6$	6.508	< 0.001	2.762	0.003	-0.119	0.547
$H = 7$	5.756	< 0.001	2.696	0.004	0.066	0.474
$H = 8$	5.142	< 0.001	3.060	0.001	0.324	0.373
$H = 9$	4.883	< 0.001	3.008	0.001	0.011	0.496
$H = 10$	4.750	< 0.001	2.871	0.002	0.231	0.408

Table 1 shows the normalized test statistic and the corresponding p -values for $H = 1, \dots, 10$ for all three panels. The El Niño anomaly typically happens at irregular intervals of three to six, sometimes seven years, thus, importance should be given to the results for these lags. However, the test rejects H_0 for the SST panel for all lags tested at any reasonable level of significance. For the South-West precipitation data, the rejection is also convincing, even

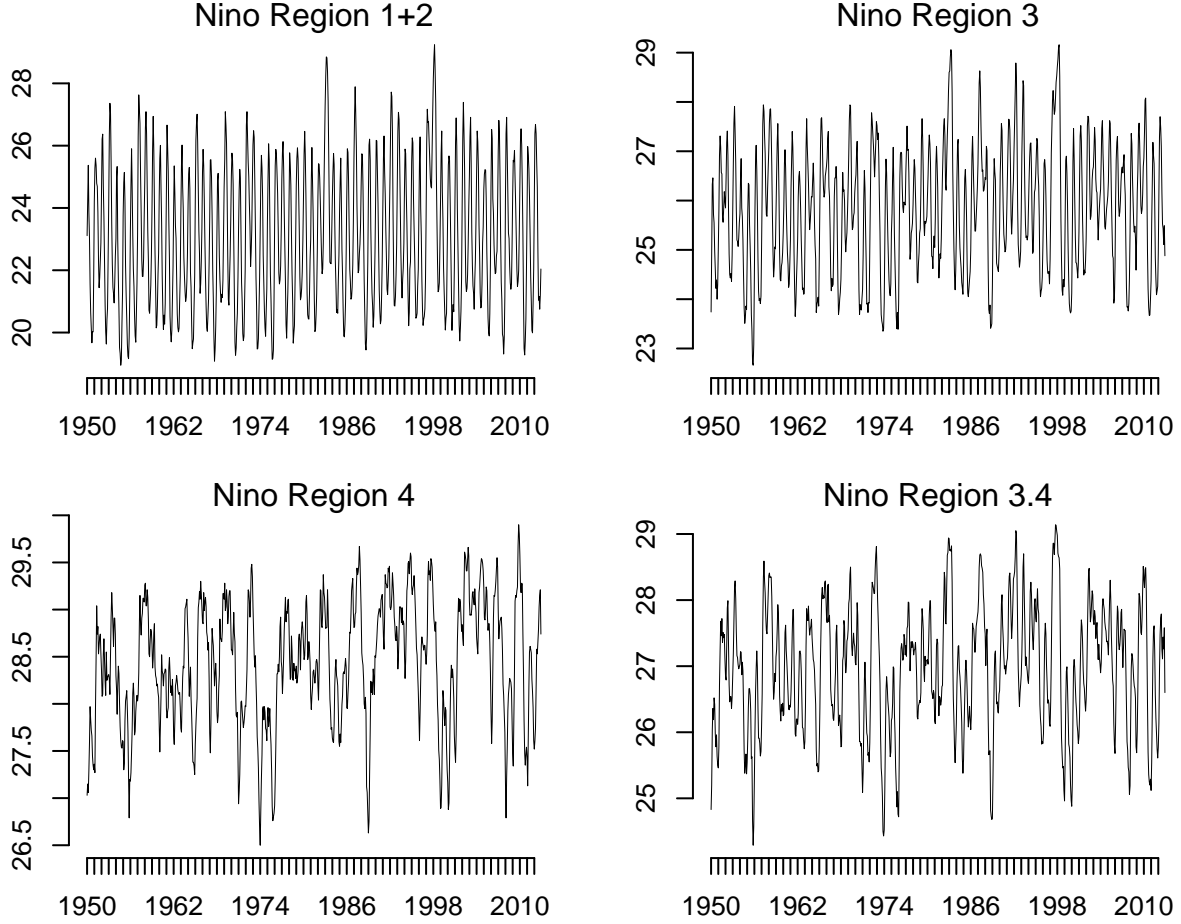


Figure 4: Sea Surface Temperature in four El Niño regions from 1950 to 2012

though less strong. Both data sets reflect known semi-periodic cycles extending over several years, and this is a likely reason for the rejections. For the German temperature data, in contrast, there is no evidence for a violation of H_0 . After a simple detrending, the panel of annual temperature curves over Germany can be taken to consist of iid observations. Due to a dense spatial coverage, one might say that after local trends have been removed, the annual temperature pattern over Germany can be treated each year as an independent replication.

3.2 Simulation scheme and finite sample performance

We now use the estimated stochastic structure of the panels described in Section 3.1 to generate artificial functional panels. This section serves a twofold purpose: we want to validate the conclusions implied by Table 1, and we want to evaluate the empirical size and power of the test in realistic settings.

3.2.1 Simulation scheme

We first describe the procedure to simulate functional panels which closely resemble the original climate data, but do not violate the null hypothesis. Then, we explain how we generate panels with increasing temporal dependence. Functional panel data are expected to exhibit correlation of curves for a fixed period n ('between' individuals, i.e., stations or regions), as well as dependence over periods for a fixed individual i ('within' a time series). The following data generating process features the first type of between correlation but excludes dependence over time: For every individual station or region i , we calculate the empirical mean function $\hat{\mu}_i(t)$ together with $k = 1, \dots, 12$ empirical principal components $\hat{v}_{k,i}(t)$ (12 is the maximum number of EFPC's for these data). We calculate the score $\hat{\xi}_{k,i,n}$ for each principal component k , for every individual i , and for every year $n = 1, \dots, N$. Let

$$\sigma_k(i, i') = \frac{1}{N-1} \sum_{n=1}^N \left(\hat{\xi}_{k,i,n} - \bar{\xi}_{k,i} \right) \left(\hat{\xi}_{k,i',n} - \bar{\xi}_{k,i'} \right)$$

and

$$\mathbf{\Sigma}_k = [\sigma_k(i, i'), 1 \leq i, i' \leq I].$$

The matrix $\mathbf{\Sigma}_k$ is the empirical covariance matrix of the scores from different individuals. Calculate its Cholesky decomposition $\mathbf{\Sigma}_k = \mathbf{L}_k \mathbf{L}_k^\top$ and obtain simulated scores of the form

$$\boldsymbol{\zeta}_k = \mathbf{z}_k \mathbf{L}_k^\top, \quad 1 \leq k \leq 12,$$

where \mathbf{z}_k is an $N \times I$ matrix of independent standard normal random variables. Each matrix $\boldsymbol{\zeta}_k$, $1 \leq k \leq 12$, has I column vectors $\boldsymbol{\zeta}_{k,i}$ of length N which are correlated among each other much like the scores of the individual stations or regions from the original climate data. With these scores, the data generating process for the simulated iid functional panel is

$$\mathbf{X}_i^{H_0}(t) = \hat{\mu}_i(t) + \sum_{k=1}^{12} \boldsymbol{\zeta}_{k,i} \hat{v}_{k,i}(t), \quad i = 1, \dots, I,$$

where each $\mathbf{X}_i^{H_0}(t)$ is a vector of random curves of length N and the superscript H_0 indicates that the artificial sample satisfies the null hypothesis of independence. The essence of the above procedure is that if the original data satisfied H_0 and were normal, then the data generating process would have the same random structure as the estimated structure of the data. Normal QQ-plots show that the scores of the three data sets in question are approximately normal (i, k fixed, N points per plot).

To construct an alternative to H_0 , we impose autocorrelation on each time series in the form of a functional autoregressive process of order 1, FAR(1) (see Chapters 3 and 4 of Bosq [2], or Chapter 13 of Horváth and Kokoszka [24]). An appropriate Cholesky factor \mathbf{L}_{ac} is defined as follows: Choose $\rho \neq 0$, $-1 < \rho < 1$, the level of autocorrelation to be imposed (one could also specify different levels of ρ for each k). Construct an $N \times N$ Toeplitz matrix

such that the first column corresponds to the sequence $\{\rho^{n-1}\}_{n=1,\dots,N}$. For the Cholesky factor \mathbf{L}_{ac} , take the lower triangular of this Toeplitz matrix and divide each element by $[(\rho^{2n} - 1)/(\rho^2 - 1)]^{1/2}$, where $n = 1, \dots, N$ denotes the row number of the corresponding element. This ensures that $\mathbf{L}_{ac}\mathbf{L}_{ac}^\top$ is positive semi-definite and has all diagonal elements equal to 1. Thus, applying this factor to a vector of length N imposes autocorrelation among the vectors' elements but does not change the overall level of variance. With the generic scores defined before, the data generating process for the autocorrelated functional panel is

$$\mathbf{X}_i^{ac}(t) = \hat{\mu}_i(t) + \sum_{k=1}^{12} \mathbf{L}_{ac} \boldsymbol{\zeta}_{k,i} \hat{v}_{k,i}(t), \quad i = 1, \dots, I.$$

In summary, we can generate siblings of functional panels of the type $\{X_{i,n}^{H_0}(t), X_{i,n}^{ac}(t)\}$, $i = 1, \dots, I$, $n = 1, \dots, N$, whose cross-sectional dependence structure is the same and similar to that of the real data, but where one sample obeys the hypothesis of independence while the other sample follows an explicit FAR(1) process. This procedure allows us to vary the length of the panel, N .

Finite sample performance

To evaluate the empirical size and power, we simulate $R = 10^3$ replications of panels with $I = 4$, $I = 42$, and $I = 103$, and with the cross-sectional dependence structure resembling that of the respective data sets. We report results for the length $N = 60$ and $N = 120$ (years), typical sample sizes encountered in historical climate data. We test for $H = 3, \dots, 6$ which is the relevant range of years for which dependence in the El Niño driven climatic measures is expected. Table 2 shows the point estimates of the rejection frequency for the simulation under H_0 together with the Clopper–Pearson confidence intervals for the probability of success (see Clopper and Pearson [9]).¹

The empirical sizes reported in Table 2 validate the results obtained in Section 3.1. For the DGPs we considered, the test has overall a satisfactory, often excellent, empirical size. The evaluation of power is more subjective as it depends on the distance of the DGP from H_0 . For every simulated panel following H_0 , we obtained an autocorrelated sibling with a fixed level of autocorrelation: $\rho = 0.38$ for $I = 4$, $\rho = 0.37$ for $I = 42$, and $\rho = 0.19$ for $I = 103$. These are the correlation levels for which the power is almost or exactly equal to 1 if $N = 120$. In light of these moderate levels of autocorrelation and the rejection frequencies reported in Table 3, we conclude that our test has excellent power together with good empirical size, such that the rejections as well as the non-rejection of H_0 for the data presented in Section 3.1 provide reliable insights.

In our simulation study, the power generally decreases with H , but it increases with H for the DGP mimicking the US precipitation panel. This agrees with the p -values reported

¹ Clopper–Pearson confidence intervals are almost identical to the confidence intervals based on the normal approximation to the binomial distribution, except for cases of empirical power close to 1 when the right end point of the latter exceeds 1.

in Table 1. The issue of the selection of H is a difficult one, and it is not satisfactorily solved even for the standard Ljung–Box–Pierce test for a single scalar time series. Statistical software packages display the p -values, often in the form of a graph, as a function of H . If for some relatively large range of H the p -values are above the 5% level, H_0 is accepted, if they are below, H_0 is rejected. In mixed cases, the test is found to be inconclusive. The same strategy can be followed in the application of our test. In addition, some background knowledge of the science problem may be utilized. In the SST and US precipitation examples, the temperature and rainfall patterns are known to reoccur every 3–6 years, so importance was attached to these lags H .

A Proof of Theorems 1 and 3

The plan of the proof of Theorem 1 is as follows. In Lemma 2, we show that the convergence to the normal limit holds if instead of projections on the EFPCs $\hat{v}_{i,j}$, projections on the $v_{i,j}$ are used. Recall that $v_{i,j}$ is the j^{th} FPC of the i^{th} functional time series in the panel. The statistic constructed using the $v_{i,j}$ is denoted by Q_N . The proof of Lemma 2 relies on Theorem 2. Next, we show in Lemmas 3 and 4 that the transition from Q_N to \hat{Q}_N involves asymptotically negligible terms. Lemma 5 collects several properties referred to in the proofs. The proof of Theorem 3 is given at the end of this section.

Let $\mathbf{X}_{i,n}$ be the column vector of length $p(i)$ defined by $\mathbf{X}_{i,n} = [\langle X_{i,n}, v_{i,j} \rangle, 1 \leq j \leq p(i)]^\top$. By stacking these I vectors on top of each other, we construct a column vector of length $p = \sum_{i=1}^I p(i)$ defined by

$$\mathbf{X}_n = [\mathbf{X}_{1,n}^\top, \dots, \mathbf{X}_{I,n}^\top]^\top.$$

We abuse notation slightly as \mathbf{X}_n was used earlier to reference the functional panel vector, but throughout this section \mathbf{X}_n will be defined as above. We allow the number of time series, I , and/or the number of FPCs, $p(i)$, used for each series to increase with the temporal sample size, N , in any way which implies that $p = p_N$ increases to infinity. Recall that $\mathbf{C}_{0,N} = \mathbb{E} [\mathbf{X}_n \mathbf{X}_n^\top]$ is the $p_N \times p_N$ covariance matrix whose eigenvalues are $\gamma_1 \geq \dots \geq \gamma_{p_N}$. Our first lemma contains two bounds involving the γ_j , which will be used throughout the proof of Theorem 1.

Lemma 1. *We have the bounds*

$$\sum_{j=1}^{p_N} \gamma_j \leq IM^{1/2} \quad \text{and} \quad p_N \leq \gamma_{p_N}^{-1} IM^{1/2},$$

where M is the bound in Assumption 1.

Proof. Notice that

$$\sum_{j=1}^{p_N} \gamma_j = \text{trace}(\mathbf{C}_{0,N}) = \mathbb{E} \|\mathbf{X}_n\|^2 \leq \sum_{i=1}^I \mathbb{E} \|X_{i,n}\|^2.$$

Applying Jensen's inequality and Assumption 1 gives the first claim. We then immediately obtain the second claim since $\gamma_{p_N} p_N \leq \sum_{j=1}^{p_N} \gamma_j$. \square

Lemma 2. *If Assumptions 1 and 2 hold then, under H_0 ,*

$$\frac{Q_N - p_N^2 H}{p_N \sqrt{2H}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Proof. The standardized vectors used in Theorem 2 are given by

$$\mathbf{Z}_{n,N} = \mathbf{C}_{0,N}^{-1/2} \mathbf{X}_n.$$

Using the mixed-product property of the Kronecker product, we can express Q_N as

$$\begin{aligned} & N^{-2} \sum_{n=1}^{N-h} \sum_{n'=1}^{N-h} (\mathbf{X}_n \otimes \mathbf{X}_{n+h})^\top (\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} (\mathbf{X}_n \otimes \mathbf{X}_{n+h}) \\ &= N^{-2} \sum_{n=1}^{N-h} \sum_{n'=1}^{N-h} (\mathbf{X}_n \otimes \mathbf{X}_{n+h})^\top (\mathbf{C}_{0,N}^{-1/2} \otimes \mathbf{C}_{0,N}^{-1/2}) (\mathbf{C}_{0,N}^{-1/2} \otimes \mathbf{C}_{0,N}^{-1/2}) (\mathbf{X}_n \otimes \mathbf{X}_{n+h}) \\ &= N^{-2} \sum_{n=1}^{N-h} \sum_{n'=1}^{N-h} (\mathbf{Z}_{n,N} \otimes \mathbf{Z}_{n+h,N})^\top (\mathbf{Z}_{n,N} \otimes \mathbf{Z}_{n+h,N}) \\ &= N^{-2} \sum_{n=1}^{N-h} \sum_{n'=1}^{N-h} |\mathbf{Z}_{n,N} \otimes \mathbf{Z}_{n+h,N}|^2. \end{aligned}$$

So we need to establish that

$$N^{-1/2} \mathbb{E} |\mathbf{Z}_{1,N}|^4 \rightarrow 0.$$

By definition we have that

$$|\mathbf{Z}_{1,N}|^4 = (\mathbf{X}_1^\top \mathbf{C}_{0,N}^{-1} \mathbf{X}_1)^2.$$

Applying the Cauchy-Schwarz and operator norm inequality we have that

$$\mathbf{X}_1^\top \mathbf{C}_{0,N}^{-1} \mathbf{X}_1 \leq |\mathbf{X}_1| \|\mathbf{C}_{0,N}^{-1} \mathbf{X}_1\| \leq |\mathbf{X}_1|^2 \|\mathbf{C}_{0,N}^{-1}\|.$$

Since $\|\mathbf{C}_{0,N}^{-1}\|$ is the largest eigenvalue of $\mathbf{C}_{0,N}^{-1}$, it is simply the reciprocal of the smallest eigenvalue of $\mathbf{C}_{0,N}$. Therefore we have

$$|\mathbf{Z}_{1,N}|^4 \leq |\mathbf{X}_1|^4 \gamma_{p_N}^{-2}.$$

The norm of \mathbf{X}_1 can be expressed as

$$|\mathbf{X}_1|^2 = \sum_i^I \sum_j^{p(i)} \langle X_{i,1}, v_{j,i} \rangle^2 \leq \sum_i^I \|X_{i,n}\|^2.$$

A final application of the Cauchy–Schwarz inequality yields

$$|\mathbf{Z}_{1,N}|^4 \leq \gamma_{p_N}^{-2} I \sum_{i=1}^I \|X_{i,1}\|^4.$$

Taking expected values we have that

$$\mathbb{E} |\mathbf{Z}_{1,N}|^4 \leq \gamma_{p_N}^{-2} I^2 M.$$

Therefore by Assumption 2

$$N^{-1/2} \mathbb{E} |\mathbf{Z}_{1,N}|^4 = N^{-1/2} \gamma_{p_N}^{-2} I^2 = o(1).$$

Finally, to apply Theorem 2 we need only to show that $p_N N^{-2/3} \rightarrow 0$. Using Lemma 1 we have that

$$p_N N^{-2/3} \leq N^{-2/3} \gamma_{p_N}^{-1} I M^{1/2}.$$

Which is clearly $o(1)$ by Assumption 2. \square

For the next lemma it will be notationally useful to define the lag h cross covariance operators:

$$\Delta_{i,i',h} = N^{-1/2} \sum_{n=1}^{N-h} X_{i,n} \otimes X_{i',n+h}.$$

Lemma 3. *If Assumptions 1 and 2 hold then, under H_0 ,*

$$\frac{Q_N - N \sum_{h=1}^H \widehat{V}_h^\top (\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} \widehat{V}_h}{p_N \sqrt{2H}} = o_P(1).$$

Proof. A minor rearrangement yields

$$\frac{Q_N - N \sum_h^H \widehat{V}_h^\top (\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} \widehat{V}_h}{p_N \sqrt{2H}} = \frac{2N \sum_h^H (V_h - \widehat{V}_h)^\top (\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} (V_h + \widehat{V}_h)}{p_N \sqrt{2H}}.$$

The Cauchy–Schwarz and operator norm inequality yield

$$|(V_h - \widehat{V}_h)^\top (\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} (V_h + \widehat{V}_h)| \leq \|V_h - \widehat{V}_h\| \|V_h + \widehat{V}_h\| \gamma_{p_N}^{-2}$$

For each coordinate of V_h , there exists i, j, i', j' such that the coordinate can be expressed as

$$N^{-1} \sum_{n=1}^{N-h} \langle X_{i,n}, v_{j,i} \rangle \langle X_{i',n+h}, v_{j',i'} \rangle = N^{-1} \sum_{n=1}^{N-h} \langle X_{i,n} \otimes X_{i',n+h}, v_{j,i} \otimes v_{j',i'} \rangle = N^{-1/2} \langle \Delta_{i,i',h}, v_{j,i} \otimes v_{j',i'} \rangle.$$

Therefore

$$\begin{aligned}
\|V_h - \widehat{V}_h\|^2 &= N^{-1} \sum_{i,i'}^I \sum_j^{p(i)} \sum_{j'}^{p(i')} \langle \Delta_{i,i',h}, v_{j,i} \otimes v_{j',i'} - \hat{v}_{j,i} \otimes \hat{v}_{j',i'} \rangle^2 \\
&\leq N^{-1} \max \|\Delta_{i,i',h}\|^2 \sum_{i,i'}^I \sum_j^{p(i)} \sum_{j'}^{p(i')} \|v_{j,i} \otimes v_{j',i'} - \hat{v}_{j,i} \otimes \hat{v}_{j',i'}\|^2 \\
&\leq N^{-1} \max \|\Delta_{i,i',h}\|^2 \sum_{i,i'}^I \sum_j^{p(i)} \sum_{j'}^{p(i')} (\alpha_{i,j} \|C_{ii} - \widehat{C}_{ii}\| + \alpha_{i',j'} \|C_{i'i'} - \widehat{C}_{i'i'}\|)^2 \\
&\leq N^{-1} \max \|\Delta_{i,i',h}\|^2 \max \|C_{ii} - \widehat{C}_{ii}\|^2 \Gamma_N \\
&= N^{-2} I^3 \Gamma_N O_P(1),
\end{aligned}$$

where the last equality follows from Lemma 5. By Parceval's inequality

$$\|V_h\|^2 = N^{-1} \sum_{i,i'}^I \sum_j^{p(i)} \sum_{j'}^{p(i')} \langle \Delta_{i,i',h}, v_{j,i} \otimes v_{j',i'} \rangle^2 \leq N^{-1} \sum_{i,i'} \|\Delta_{i,i',h}\|^2 = N^{-1} O_P(I^2),$$

and the same holds for $\|\widehat{V}_h\|^2$. Combining everything, the original difference is of the order

$$N p_N^{-1} \gamma_{p_N}^{-2} N^{-1} I^{3/2} \Gamma_N^{1/2} N^{-1/2} I O_P(1) = N^{-1/2} p_N^{-1} I^{5/2} \gamma_{p_N}^{-2} \Gamma_N^{1/2} O_P(1),$$

which is $o_P(1)$ by Assumption 2. □

Lemma 4. *If Assumptions 1 and 2 hold, then under H_0 ,*

$$\frac{N \sum_{h=1}^H [\widehat{V}_h^\top (\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} \widehat{V}_h - \widehat{V}_h^\top (\widehat{\mathbf{C}}_{0,N} \otimes \widehat{\mathbf{C}}_{0,N})^{-1} \widehat{V}_h]}{p_N \sqrt{2H}} = o_P(1).$$

Proof. Using the Cauchy–Schwarz inequality and operator norm inequality we have that

$$\begin{aligned}
&|\widehat{V}_h^\top (\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} \widehat{V}_h - \widehat{V}_h^\top (\widehat{\mathbf{C}}_{0,N} \otimes \widehat{\mathbf{C}}_{0,N})^{-1} \widehat{V}_h| \\
&\leq \|\widehat{V}_h^\top\|^2 \|(\mathbf{C}_{0,N} \otimes \mathbf{C}_{0,N})^{-1} - (\widehat{\mathbf{C}}_{0,N} \otimes \widehat{\mathbf{C}}_{0,N})^{-1}\|_S \\
&\leq \|\widehat{V}_h\|^2 [\gamma_{p_N}^{-1} + \widehat{\gamma}_{p_N}^{-1}] \|\mathbf{C}_{0,N}^{-1} - \widehat{\mathbf{C}}_{0,N}^{-1}\|_S \\
&\leq \|\widehat{V}_h\|^2 [\gamma_{p_N}^{-1} + \widehat{\gamma}_{p_N}^{-1}] [\gamma_{p_N}^{-1} \widehat{\gamma}_{p_N}^{-1}] \|\mathbf{C}_{0,N} - \widehat{\mathbf{C}}_{0,N}\|_S \\
&\leq \|\widehat{V}_h\|^2 [\gamma_{p_N}^{-1} + \widehat{\gamma}_{p_N}^{-1}] [\gamma_{p_N}^{-1} \widehat{\gamma}_{p_N}^{-1}] \|\mathbf{C}_{0,N} - \widehat{\mathbf{C}}_{0,N}\|.
\end{aligned}$$

As before, we can apply Lemma 5.1 to obtain

$$\|\widehat{V}_h\|^2 = \widehat{V}_h^\top \widehat{V}_h = N^{-1} \sum_{i,i'}^I \sum_j^{p(i)} \sum_{j'}^{p(i')} \langle \Delta_{i,i',h}, \hat{v}_{j,i} \otimes \hat{v}_{j',i'} \rangle^2 \leq N^{-1} \sum_{i,i'} \|\Delta_{i,i',h}\|^2 = O_P(N^{-1} I^2).$$

Turning to the eigenvalues we have that

$$[\gamma_{p_N}^{-1} + \hat{\gamma}_{p_N}^{-1}][\gamma_{p_N}^{-1} \hat{\gamma}_{p_N}^{-1}] = \gamma_{p_N}^{-3} \left[\frac{\gamma_{p_N}}{\hat{\gamma}_{p_N}} + \frac{\gamma_{p_N}^2}{\hat{\gamma}_{p_N}^2} \right].$$

Applying Lemma 5.4 to bound the difference

$$\left| \frac{\hat{\gamma}_{p_N}}{\gamma_{p_N}} - 1 \right| = \frac{|\hat{\gamma}_{p_N} - \gamma_{p_N}|}{\gamma_{p_N}} \leq \frac{\|\hat{\mathbf{C}}_{0,N} - \mathbf{C}_{0,N}\|}{\gamma_{p_N}} = O_P(\gamma_{p_N}^{-1} I N^{-1/2} \Gamma_N^{1/2}) = o_P(1).$$

Putting everything together, the original difference is

$$O_P(\gamma_{p_N}^{-3} I^3 N^{-1/2} p_N^{-1} \Gamma_N^{1/2}) = o_P(1),$$

by Assumption 2. □

Our last lemma contains several properties which were used in the arguments developed above.

Lemma 5. *If Assumptions 1 and 2 hold, then we have the following properties:*

1. $\max \|\Delta_{i,i',h}\|^2 = O_P(I^2)$ under H_0 .
2. $\max \|C_{ii'}\| = O(1)$.
3. $\max \|\hat{C}_{ii} - C_{ii}\| = O_P(N^{-1/2} I)$ and $\max \|\hat{C}_{ii} - C_{ii}\|^2 = O_P(N^{-1} I)$
4. $\|\hat{\mathbf{C}}_{0,N} - \mathbf{C}_{0,N}\| = O_P(IN^{-1/2} \Gamma_N^{1/2})$

Proof. 1. For each fixed i and i' , we have that

$$\begin{aligned} \mathbb{E} \|\Delta_{i,i',h}\|^2 &= N^{-1} \sum_{n=1}^{N-h} \sum_{n'=1}^{N-h} \langle X_{i,n} \otimes X_{i,n+h}, X_{i',n'} \otimes X_{i',n'+h} \rangle \\ &= N^{-1} \sum_{n=1}^{N-h} \mathbb{E} \langle X_{i,n} \otimes X_{i,n+h}, X_{i',n} \otimes X_{i',n+h} \rangle \\ &\leq \mathbb{E} \|X_{i,1}\|^2 \mathbb{E} \|X_{i',1}\|^2 \leq M. \end{aligned}$$

Therefore we have that

$$\mathbb{E} [\max \|\Delta_{i,i',h}\|^2] \leq I^2 M,$$

and the result follows from Markov's inequality.

2. By Jensen's inequality we have that

$$\begin{aligned} \|C_{ii'}\| &\leq \mathbb{E} \|X_{i,n} \otimes X_{i',n}\| = \mathbb{E} [\|X_{i,n}\| \|X_{i',n}\|] \leq \sqrt{\mathbb{E} \|X_{i,n}\|^2 \mathbb{E} \|X_{i',n}\|^2} \\ &\leq (\mathbb{E} \|X_{i,n}\|^4 \mathbb{E} \|X_{i',n}\|^4)^{1/4} \leq M^{1/2}, \end{aligned}$$

which proves the claim.

3. The argument is the same as in 1.
4. By the triangle inequality we have

$$\|\widehat{\mathbf{C}}_{0,N} - \mathbf{C}_{0,N}\| \leq \|\widehat{\mathbf{C}}_{0,N} - \tilde{\mathbf{C}}_{0,N}\| + \|\tilde{\mathbf{C}}_{0,N} - \mathbf{C}_{0,N}\|$$

where $\tilde{\mathbf{C}}_{0,N}$ is formed by projecting the $C_{ii'}$ onto the estimated PCs. So the square of the first term is given by

$$\begin{aligned} \|\widehat{\mathbf{C}}_{0,N} - \tilde{\mathbf{C}}_{0,N}\|^2 &= \sum_{i,i'}^I \sum_{j=1}^{p(i)} \sum_{j'=1}^{p(i')} \left(N^{-1} \sum_{n=1}^N \langle X_{i,n} \otimes X_{i',n}, \hat{v}_{j,i} \otimes \hat{v}_{j',i'} \rangle - \langle C_{i,i'}, \hat{v}_{j,i} \otimes \hat{v}_{j',i'} \rangle \right)^2 \\ &\leq \sum_{i,i'}^I \left\| N^{-1} \sum_{n=1}^N X_{i,n} \otimes X_{i',n} - C_{i,i'} \right\|^2 = O_P(N^{-1}I^2). \end{aligned}$$

The square of the second term is given by

$$\begin{aligned} \|\tilde{\mathbf{C}}_{0,N} - \mathbf{C}_{0,N}\|^2 &= \sum_{i,i'}^I \sum_{j=1}^{p(i)} \sum_{j'=1}^{p(i')} \langle C_{i,i'}, \hat{v}_{j,i} \otimes \hat{v}_{j',i'} - v_{j,i} \otimes v_{j',i'} \rangle^2 \\ &\leq \sum_{i,i'}^I \sum_{j=1}^{p(i)} \sum_{j'=1}^{p(i')} \|C_{i,i'}\|^2 (\|v_{j,i} - \hat{v}_{j,i}\| + \|v_{j',i'} - \hat{v}_{j',i'}\|)^2 \\ &\leq \sum_{i,i'}^I \sum_{j=1}^{p(i)} \sum_{j'=1}^{p(i')} \|C_{i,i'}\|^2 (\sqrt{2}\alpha_{j,i}^{-1} \|C_{ii} - \widehat{C}_{ii}\| + \sqrt{2}\alpha_{j',i'}^{-1} \|C_{i'i'} - \widehat{C}_{i'i'}\|)^2 \\ &\leq 2 \max_{i,i'} \|C_{i,i'}\|^2 \max_{ii} \|C_{ii} - \widehat{C}_{ii}\|^2 \sum_{i,i'}^I \sum_{j=1}^{p(i)} \sum_{j'=1}^{p(i')} (\alpha_{j,i}^{-1} + \alpha_{j',i'}^{-1})^2 \\ &= IN^{-1}\Gamma_N O_P(1). \end{aligned}$$

Therefore both terms are asymptotically bounded by $I^2 N^{-1} \Gamma_N O_P(1)$, which proves the claim. \square

Proof of Theorem 3. Analogous results to Lemmas 3 and 4 are obtained in the same way and are thus omitted for brevity. We mention that the key difference is that the $\|\Delta_{i,i',h}\|$ is no longer of order $O_P(1)$, but $O_P(N^{1/2})$. The \widehat{C}_{ii} are still root- N consistent since the series is assumed to be L^4 -m approximable. We therefore only show that

$$\frac{Q_N - p_N^2 H}{p_N \sqrt{2H}} \xrightarrow{P} \infty.$$

We assume that those lag terms which exhibit correlation are contained in the set \mathcal{H}^* we therefore begin with the lower bound

$$Q_N \geq N \sum_{h \in \mathcal{H}^*} \mathbf{V}_{h^*}^\top (\mathbf{C}_0 \otimes \mathbf{C}_0)^{-1} \mathbf{V}_{h^*}.$$

Since the smallest eigenvalue of $(\mathbf{C}_0 \otimes \mathbf{C}_0)^{-1}$ is γ_1^{-2} we can bound Q_N below using

$$Q_N \geq N\gamma_1^{-2} \sum_{h \in \mathcal{H}^*} \|\mathbf{V}_h\|^2.$$

Isolating the pairs \mathcal{I}_h^* which are assumed to be correlated (at a lag of h) we can further bound below as

$$Q_N \geq \gamma_1^{-2} N \sum_{h \in \mathcal{H}^*} \sum_{(i,j) \in \mathcal{I}_h^*} \left(N^{-1} \sum_{n=1}^N (\langle X_n, v_i \rangle \langle X_{n+h}, v_j \rangle) \right)^2 = \gamma_1^{-2} NR(1 + o_P(1)) \sum_{h \in \mathcal{H}^*} |\mathcal{I}_h|$$

where the last equality holds since, by Assumption 3, the summands form a stationary and ergodic sequence. Combining Lemma 1 with Assumption 2, $N\gamma_1^{-2}$ tends to infinity faster than p_N^2 and the claim holds. We also see that the effect of having more indices which exhibit correlation is additive. □

References

- [1] Aston, J. A. D. and Kirch, C. (2012). Evaluating stationarity via change-point alternatives with application to fMRI data. *The Annals of Applied Statistics*, 6:1906–1948.
- [2] Bosq, D. (2000). *Linear Processes in Function Spaces*. Springer.
- [3] Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive moving average time series models. *Journal of the American Statistical Association*, 65:1509–1526.
- [4] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer, New York.
- [5] Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer, second edition.
- [6] Cai, T. and Hall, P. (2006). Prediction in functional linear regression. *The Annals of Statistics*, 34:2159–2179.
- [7] Cardot, H., Ferraty, F., Mas, A., and Sarda, P. (2003). Testing hypothesis in the functional linear model. *Scandinavian Journal of Statistics*, 30:241–255.
- [8] Chitturi, R. V. (1976). Distribution of multivariate white noise autocorrelation. *Journal of the American Statistical Association*, 71:223–226.
- [9] Clopper, C. J. and Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26:404 – 413.
- [10] Crambes, C., Kneip, A., and Sarda, P. (2009). Smoothing splines estimators for functional linear regression. *The Annals of Statistics*, 37:35–72.
- [11] Dauxois, J., Pousse, A., and Romain, Y. (1982). Asymptotic theory for principal component analysis of a vector random function. *Journal of Multivariate Analysis*, 12:136–154.

- [12] Fisher, T. J. and Gallagher, C. M. (2012). New weighted portmanteau statistics for time series goodness of fit testing. *Journal of the American Statistical Association*, 107:777–787.
- [13] Fremdt, S., Horváth, L., Kokoszka, P., and Steinebach, J. (2014). Functional data analysis with increasing number of projections. *Journal of Multivariate Analysis*, 124:313–332.
- [14] Fu, B., Li, W.-K., and Fung, W.-K. (2002). Testing model adequacy for dynamic panel data with intercorrelation. *Biometrika*, 89:591–601.
- [15] Gabrys, R. and Kokoszka, P. (2007). Portmanteau test of independence for functional observations. *Journal of the American Statistical Association*, 102:1338–1348.
- [16] Graham, A. (1981). *Kronecker Products and Matrix Calculus with Applications*. John Wiley and Sons.
- [17] Hall, P., Müller, H.-G., and Wang, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *The Annals of Statistics*, 34:1493–1517.
- [18] Härdle, W. and Osipenko, M. (2012). Spatial risk premium on weather derivatives and hedging weather exposure in electricity. *The Energy Journal*, 33:149–170.
- [19] Härdle, W. K. and Majer, P. (2015). Yield curve modeling and forecasting using semiparametric factor dynamics. *The European Journal of Finance*, 00:DOI: 10.1080/1351847X.2014.926281. Forthcoming.
- [20] Hjellvik, V. and Tjøstheim, D. (1999). Modelling panels of intercorrelated autoregressive time series. *Biometrika*, 86:573–590.
- [21] Hjort, N. L., Mckeague, I. W., and Van Keilegom, I. (2009). Extending the scope of empirical likelihood. *The Annals of Statistics*, 37:1079–1111.
- [22] Hörmann, S., Kidziński, L., and Hallin, M. (2015). Dynamic functional principal components. *Journal of the Royal Statistical Society: Series B*, 77(2):319–348.
- [23] Horváth, L., Hušková, M., and Rice, G. (2013). Test of independence for functional data. *Journal of Multivariate Analysis*, 117:100–119.
- [24] Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer.
- [25] Hosking, J. R. M. (1980). The multivariate portmanteau statistic. *Journal of the American Statistical Association*, 75:602–608.
- [26] Hsiao, C. (2003). *Analysis of Panel Data*. Econometric Society.
- [27] Jiofack, J. G. A. and Nkiet, G. M. (2010). Testing for lack of dependence between functional variables. *Statistics and Probability Letters*, 80:1210–1217.
- [28] Jirak, M. (2015). Optimal eigen expansions and uniform bounds. *Probability Theory and Related Fields*, 00:1–47. Forthcoming, DOI 10.1007/s00440-015-0671-3.
- [29] Kowal, D. R., Matteson, D. S., and Ruppert, D. (2016). A bayesian multivariate functional dynamic linear model. *Journal of the American Statistical Association*. Forthcoming.
- [30] Li, B. and Song, J. (2016). Nonlinear sufficient dimension reduction for functional data. *The Annals of Statistics*. Forthcoming.

- [31] Li, W. K. (2004). *Diagnostic Checks in Time Series*. Chapman and Hall.
- [32] Liebl, D. (2013). Modeling and forecasting electricity spot prices: A functional data perspective. *The Annals of Applied Statistics*, 7:1249–1835.
- [33] Ljung, G. and Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, 66:67–72.
- [34] Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.
- [35] McLeod, A. I. (1978). On the distribution of residual autocorrelations in Box–Jenkins models. *Journal of the Royal Statistical Society (B)*, 40:296–302.
- [36] Panaretos, V. and Tavakoli, S. (2013). Cramér–karhunen–loève representation and harmonic principal component analysis of functional time series. *Stochastic Processes and their Applications*, 123(7):2779–2807.
- [37] Paul, D. and Peng, J. (2009). Consistency of restricted maximum likelihood estimators of principal components. *The Annals of Statistics*, 37:1229–1271.
- [38] Peng, H. and Schick, A. (2012). Asymptotic normality of quadratic forms with random vectors of increasing dimension. Technical report, Binghamton University. <ftp://ftp.math.binghamton.edu/pub/anton/cltqf.pdf>.
- [39] Peng, H. and Schick, A. (2013). Empirical likelihood approach to goodness of fit testing. *Bernoulli*, 13:954–981.
- [40] Peña, D. and Rodriguez, J. (2002). A powerful portmanteau test of lack of fit for time series. *Journal of the American Statistical Association*, 97:601–610.
- [41] Ramsay, J., Hooker, G., and Graves, S. (2009). *Functional Data Analysis with R and MATLAB*. Springer.
- [42] Senatov, V. V. (1998). *Normal Approximation: New Results, Methods and Problems*. VSP.
- [43] Trenberth, K. E. (1997). The definition of El Niño. *Bulletin of the American Meteorological Society*, 78:2771–2777.
- [44] Trenberth, K. E. and Stepaniak, D. P. (2001). Indices of El Niño evolution. *Journal of Climate*, 14:1697–1701.
- [45] Zhang, X., Shao, X., Hayhoe, K., and Wuebbles, D. (2011). Testing the structural stability of temporally dependent functional observations and application to climate projections. *Electronic Journal of Statistics*, 5:1765–1796.

Table 2: Rejection frequencies (and confidence bands) for the test with $\alpha = 0.05$ obtained from 1000 simulations under H_0 for each of the data sets.

El Niño data, $I = 4$, i.i.d.				
	$N = 60$		$N = 120$	
$H = 3$	0.055	(0.042, 0.071)	0.053	(0.040, 0.069)
$H = 4$	0.071	(0.056, 0.089)	0.051	(0.038, 0.067)
$H = 5$	0.060	(0.046, 0.077)	0.061	(0.047, 0.078)
$H = 6$	0.063	(0.049, 0.080)	0.057	(0.043, 0.073)

US precipitation data, $I = 103$, i.i.d.				
	$N = 60$		$N = 120$	
$H = 3$	0.033	(0.023, 0.046)	0.033	(0.023, 0.046)
$H = 4$	0.046	(0.034, 0.061)	0.041	(0.030, 0.055)
$H = 5$	0.054	(0.041, 0.070)	0.048	(0.036, 0.063)
$H = 6$	0.079	(0.063, 0.097)	0.058	(0.044, 0.074)

German temperature data, $I = 42$, i.i.d.				
	$N = 60$		$N = 120$	
$H = 3$	0.062	(0.048, 0.079)	0.061	(0.047, 0.078)
$H = 4$	0.060	(0.046, 0.077)	0.052	(0.039, 0.068)
$H = 5$	0.063	(0.049, 0.080)	0.054	(0.041, 0.070)
$H = 6$	0.059	(0.045, 0.075)	0.054	(0.041, 0.070)

Table 3: Rejection frequencies (and confidence bands) for the test with $\alpha = 0.05$ obtained from 1000 simulations of an AR(1) process for each of the data sets; ρ indicates the degree of autocorrelation.

El Niño SST curves, $I = 4$, $\rho = 0.38$				
	$N = 60$		$N = 120$	
$H = 3$	0.987	(0.978, 0.993)	1.000	(0.996, 1.000)
$H = 4$	0.927	(0.909, 0.942)	1.000	(0.996, 1.000)
$H = 5$	0.778	(0.751, 0.803)	1.000	(0.996, 1.000)
$H = 6$	0.607	(0.576, 0.637)	0.997	(0.991, 0.999)

US precipitation data, $I = 103$, $\rho = 0.19$				
	$N = 60$		$N = 120$	
$H = 3$	0.951	(0.936, 0.964)	1.000	(0.996, 1.000)
$H = 4$	0.981	(0.970, 0.989)	1.000	(0.996, 1.000)
$H = 5$	0.994	(0.987, 0.998)	1.000	(0.996, 1.000)
$H = 6$	0.994	(0.987, 0.998)	1.000	(0.996, 1.000)

German temperature curves, $I = 42$, $\rho = 0.37$				
	$N = 60$		$N = 120$	
$H = 3$	0.790	(0.763, 0.815)	0.996	(0.990, 0.999)
$H = 4$	0.690	(0.660, 0.719)	0.986	(0.977, 0.992)
$H = 5$	0.615	(0.584, 0.645)	0.974	(0.962, 0.983)
$H = 6$	0.554	(0.523, 0.585)	0.966	(0.953, 0.976)